Relationship between the Pittsburgh Sleep Quality Index and the Epworth Sleepiness Scale in a sleep laboratory referral population

Protsanta Mondal¹
John A Gjevre²
Regina M Taylor-Gjevre³
Hyun J Lim¹
¹Department of Community Health and Epidemiology, ²Division of Respiratory, Critical Care and Sleep Medicine, ³Department of Medicine, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

Background: Sleep health questionnaires are often employed as a first assessment step for sleep pathology. The Epworth Sleepiness Scale (ESS) and the Pittsburgh Sleep Quality Index (PSQI) are two commonly employed questionnaire instruments. Aspects of sleep health may be measured differently depending on choice of instrument.

Objectives: In a patient population at high risk for sleep disorders, referred for polysomnography (PSG), we evaluated the level of association between results from these two instruments. Questionnaire results were also compared with measured PSG parameters.

Methods: Records of patients undergoing overnight PSG in the sleep laboratory between February–June 2011 were retrospectively reviewed for eligibility. Inclusion criteria were met by 236 patients. PSQI and ESS scores, demographic information, and PSG data were extracted from each record for analysis. Four subgroups based on normal/abnormal values for ESS and PSQI were evaluated for between-group differences.

Results: Of 236 adult participants, 72.5% were male, the mean age was 52.9 years (13.9), mean body mass index (BMI) 34.4 kg/m² (8.3), mean ESS 9.0 (4.8; range: 0–22), PSQI mean 8.6 (4.2; range: 2–19). The Pearson correlation coefficient was r = 0.13 (P = 0.05) for association between ESS and PSQI. Participants with an abnormal ESS were more likely to have an abnormal PSQI score (odds ratio 1.9 [1.1–3.6]; P = 0.03). Those with an abnormal ESS had higher BMI (P = 0.008) and higher apnea–hypopnea indexes (AHI) (P = 0.05). Differences between the four subgroups were observed for BMI and sex proportions, but not for AHI.

Conclusions: We observed limited association between these two commonly used questionnaire instruments, the ESS and the PSQI. These two questionnaires appear to evaluate different aspects of sleep. In terms of clinical application, for global assessment of patients with sleep problems, care should be taken to include instruments measuring different facets of sleep health.

Keywords: Pittsburgh Sleep Quality Index, Epworth Sleepiness Scale, polysomnography, sleep disorders

Introduction

Primary sleep disorders are an increasing public health concern and are definitively diagnosed by polysomnography (PSG). However, as access to PSG may be limited, sleep health questionnaires are often employed as a first assessment step.¹

Two commonly employed sleep health assessment questionnaires are the Epworth Sleepiness Scale (ESS) and the Pittsburgh Sleep Quality Index (PSQI).²³ These instruments are utilized as general measures of sleep health and daytime dysfunction, and are not specific for any single primary sleep disorder. Both the ESS and the PSQI have been widely translated and employed in a wide range of both population-based and
clinical studies. However, there have been few studies examining the relationship between the two instrument measures.

An evaluation in a general adult community population by Buysse et al demonstrated little association between the ESS and PSQI. The PSQI score was also not found to be associated with specific PSG measures. Patient subgroups were examined based on combinations of normal/abnormal PSQI and ESS scores for association with objective measures. In this general population, these questionnaire aggregates did not strongly associate with PSG findings.

In this present study we apply these questionnaires, the ESS and PSQI, to a population referred for PSG, with a high probability of sleep pathology. Our objectives were to compare PSQI and ESS scores in this high-risk population and evaluate the level of association between these two questionnaire instrument results. We hypothesize the level of association between these questionnaires would be high in such a sleep laboratory referral population. Further, we evaluate whether PSQI/ESS score aggregate-based subgroups in this referral population would exhibit differences in objective PSG characteristics.

Methods

Records of 367 consecutive patients studied by PSG in a university sleep disorders laboratory between February 2011 and June 2011 were retrospectively reviewed for inclusion in this study. Referrals to the sleep laboratory were restricted to select physicians trained in sleep medicine. All patients attending the sleep laboratory were requested to complete standard questionnaires, which include the PSQI, and the ESS. A total of 236 patients met the study inclusion criteria which included age 18 years or older, and appropriate completion of the PSQI and ESS questionnaires. Charts were excluded if the ESS or PSQI questionnaires were incomplete.

The PSQI is a seven domain (19 item) self-rated questionnaire evaluating usual sleep habits during the last month. The seven domain scores including: subjective sleep quality, sleep latency, sleep duration, sleep efficiency, daytime dysfunction, sleep fragmentation, and use of sleep aid medications; combine to provide a global sleep quality index score. The possible scores range from 0–21, with greater than five indicative of impaired sleep quality.

The ESS is an eight-item assessment of somnolence, with possible scores from 0 to 24. A score of over 10 is considered to be abnormal.

The PSQI and ESS have been demonstrated to be stable measures over time in early middle-aged adults, and have been recently validated in older men. The sleep laboratory charts were reviewed by a single researcher (PM), using a standardized data collection form. Questionnaire scores for the PSQI and ESS were extracted from the charts, as well as demographic data including: age, sex, and body mass index (BMI). PSG objective measures including: apnea hypopnea index (AHI), apnea index (AI), respiratory effort-related arousals (RERA), and sleep efficiency (SE: time asleep/total time in bed) were also recorded for each patient.

Patients were studied overnight in the sleep lab using the standard 15 channel PSG (version 8.0, sleep diagnostic software; Sandman Elite, Ottawa, Canada). Established protocols were used for all PSG studies. This included electroencephalogram (EEG, 3 channel), electrooculogram (2 channels), electromyogram (chin and leg), electrocardiogram, heart rate, snoring, thermistor airflow, nasal pressure airflow, oxygen saturation, chest wall motion, and abdomen motion.

Institutional ethics approval was obtained for this study, which complied with the Declaration of Helsinki.

Statistical methods

We checked variables regarding assumptions underlying the use of parametric and nonparametric statistics, and analyzed accordingly. Descriptive statistics were used to characterize global PSQI scores, ESS scores, demographics, and PSG measures of sleep continuity. Pearson correlation coefficients were calculated to examine the correlation between quantitative variables. We performed Student’s t-test, analyses of variance (ANOVA) for normally distributed variables (BMI, age, ESS, SE, and PSQI), Wilcoxon and Kruskal-Wallis tests for non-normally distributed variables (AHI, AI, arousing periodic limb movements (APLM), and RERA). Chi-square test was used to examine the association between categorical variables. Cronbach’s alpha was utilized to evaluate internal questionnaire consistency in the study population.

Published cut-off scores (>5 on the PSQI, >10 on the ESS) for each instrument were used to create four subgroups that had higher scores on each, normal scores on each, or one higher, and one normal score. ANOVA and Kruskal-Wallis test were conducted to find differences in means/medians among these subgroups for demographics and PSG measures. Chi-square test was used to find association between these subgroups and sex, BMI category. Data were analyzed using SAS software (v 9.3; SAS Cary, NC, USA). The level of significance was set at 0.05 (two-tailed).
Results

Measured parameters within the study population are detailed in Table 1. Of the 236 patients, the male:female proportions were 72.5% (n = 171): 27.5% (n = 65). There were no statistically significant differences in age (P = 0.32) or BMI (P = 0.30) between sex groups.

As an assessment of internal consistency, Cronbach’s alpha equaled 0.83 and 0.74 for the ESS and PSQI instruments respectively in this study sample. This represents good-to-acceptable levels of internal consistency.12

Associations were examined between the two sleep questionnaire scores. Correlations were weak with Pearson correlation coefficient (r = 0.13; P = 0.05) between the ESS and PSQI scores. ESS and PSQI raw scores in this population are illustrated by scatterplot (Figure 1). If the daytime dysfunction domain of the PSQI is not considered in the collated score, the association with the ESS is further weakened (r = 0.05; P = 0.43). In isolation, the daytime dysfunction scale score correlates moderately well with the ESS (r = 0.42; P < 0.0001). Additionally, associations were examined between the ESS, PSQI, and PSG measures. The AHI is particularly relevant in the diagnosis of obstructive sleep apnea (OSA). Between AHI and PSQI (r = −0.12; P = 0.07) and between AHI and ESS (r = 0.13; P = 0.046).

Comparison by category, normal/abnormal as illustrated in Table 2, provides an odds ratio of 1.9 (P = 0.03) for having an abnormal PSQI with an abnormal ESS.

Table 1 Measured parameters within the study population

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI (kg/m²)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>34.4</td>
<td>8.3</td>
<td>32.8</td>
<td>18.0</td>
<td>66.0</td>
</tr>
<tr>
<td>Men</td>
<td>34.1</td>
<td>7.7</td>
<td>32.5</td>
<td>18.4</td>
<td>66.0</td>
</tr>
<tr>
<td>Women</td>
<td>35.4</td>
<td>9.7</td>
<td>34.5</td>
<td>18.0</td>
<td>59.0</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>52.9</td>
<td>13.9</td>
<td>53.0</td>
<td>19.0</td>
<td>90.3</td>
</tr>
<tr>
<td>Men</td>
<td>52.3</td>
<td>13.9</td>
<td>52.4</td>
<td>19.0</td>
<td>90.3</td>
</tr>
<tr>
<td>Women</td>
<td>54.3</td>
<td>13.9</td>
<td>56.7</td>
<td>23.6</td>
<td>82.1</td>
</tr>
<tr>
<td>ESS</td>
<td>9.0</td>
<td>4.8</td>
<td>9.0</td>
<td>0</td>
<td>22.0</td>
</tr>
<tr>
<td>PSQI</td>
<td>8.6</td>
<td>4.2</td>
<td>8.0</td>
<td>2.0</td>
<td>19.0</td>
</tr>
<tr>
<td>SE%</td>
<td>68.3</td>
<td>21.1</td>
<td>74.5</td>
<td>6.4</td>
<td>97.3</td>
</tr>
<tr>
<td>AHI</td>
<td>31.3</td>
<td>36.5</td>
<td>15.5</td>
<td>0</td>
<td>158.8</td>
</tr>
<tr>
<td>APLM/hour</td>
<td>17.0</td>
<td>12.3</td>
<td>14.0</td>
<td>1.0</td>
<td>72.0</td>
</tr>
<tr>
<td>RERA/hour</td>
<td>0.5</td>
<td>1.3</td>
<td>0</td>
<td>0</td>
<td>9.4</td>
</tr>
</tbody>
</table>

Abbreviations: AHI, Apnea–Hypopnea Index; AI, Arousal Index; APLM, Arousing Periodic Limb Movements; BMI, Body Mass Index kg/m²; ESS, Epworth Sleepiness Scale; PSQI, Pittsburgh Sleep Quality Index; RERA, respiratory effort-related arousals; SD, standard deviation; SE, sleep efficiency.

Results

Of the female participants, 43.1% (28) had abnormal ESS > 10, and 86.1% (56) had abnormal PSQI > 5. In the male participants, 36.3% (62) had abnormal ESS and 64.9% (111) had abnormal PSQI scores. We examined the populations based on separate categorization by normal/abnormal PSQI and normal/abnormal ESS (Tables 3 and 4). Covariate comparison between groups reveals no statistically significant differences in PSG measures between groups based on PSQI category. Comparison between groups based on ESS status demonstrates statistically significant differences in BMI, AHI, and APLM.

We also considered potential differences between cluster groups for ESS and PSQI results. We divided the population into four groups: normal PSQI/normal ESS; normal PSQI/abnormal ESS; abnormal PSQI/normal ESS; and abnormal PSQI and abnormal ESS. Examining the population by these subgroups (Table 5), we observe statistically significant differences between groups for sex proportion, and mean BMI. There were no statistically significant differences observed between these subgroups for most PSG measured parameters.

Discussion

In this retrospective chart review study of patients suspected of sleep disorder and referred for diagnostic polysomnographic testing, we found only weak correlation between ESS and PSQI scores.

The ESS gauges a respondent’s self-reported probability of falling asleep in eight common situations. These everyday situations include while sitting and reading, watching television, sitting inactive in a public place, as a passenger in a car for an hour, while lying down in the afternoon, while sitting and talking to another, sitting quietly after lunch, and in a car while stopped in traffic.7 While the ESS score reflects subjective self-reported levels of sleepiness, an objective measure of sleepiness or hypersomnolence would be documentation of sleep latency by multiple sleep latency test (MSLT).13

Previous investigators have reported correlations between ESS and MSLT-defined sleepiness with coefficients of between 0.03 and 0.86.14-16 Using survival analysis in a large population, Aurora et al have recently reported a clear association between the ESS score and the average sleep latency as objectively measured from MSLT diagnostic testing.16 Interestingly, an ESS score of ≥13 was identified by Aurora as most effectively predicting objective sleepiness,16 in contrast to the conventional threshold score of >10.2
this subjective measure of sleepiness correlated only weakly with the PSQI ($r = 0.13; P = 0.05$). Examination of dichotomous categories of ESS and PSQI revealed an odds ratio of $1.9$ ($P = 0.03$) for an individual with an abnormal ($>10$) ESS having an abnormal PSQI ($>5$). As obstructive sleep apnea was the most common suspected diagnosis in this referral population, the ESS association with AHI was examined, with an equally weak association observed ($r = 0.13; P = 0.046$). Participants with an ESS $>10$ had statistically significantly higher BMIs and apnea hypopnea indexes consistent with OSA. These findings are consistent with those recently reported by Sil and Barr in a UK study of sleep laboratory patients.\(^{17}\)

The PSQI, in contrast to the ESS, evaluates subjective sleep quality over the previous month. The multidimensional sleep quality assessment predominantly gauges perceived nocturnal sleep quality with one component to address daytime dysfunction.\(^{3}\) In this study population, we did observe that omitting the daytime dysfunction component weakened the correlation with the ESS scores ($r = 0.05; P = 0.43$). Previous investigators have reported the strong association of PSQI scores to depression and psychosocial stress.\(^{18}\) The ESS has also been reported to have some association with depression instrument scores.\(^{19}\) Benitez and Gunstad recently reported an independent relationship between PSQI domain scores and cognitive decrements impacting attention and executive functions.\(^{20}\) It has been recognized that the PSQI may be influenced by various internal and also external factors, and may not always be associated with polysomnographic findings of a distinct primary sleep disorder such as abnormal AHI or periodic limb movements (PLMs).\(^{21,22}\) Kezirian et al found PSQI scores to be unassociated with measures of sleep disordered breathing severity in older men.\(^{21}\) Neau et al recently reported lack of correlation in multiple sclerosis patients, between PSQI scores and results of objective assessments of vigilance and sleep.\(^{23}\) In a nonclinical sample, PSQI scores correlated with depression scale results, but not with actigraphic measures of sleep.\(^{24}\)

In this current study, the mean PSQI score was 8.6 (range: 2–19). As previously mentioned, the correlation between the global PSQI score (including the daytime dysfunction component) and the ESS was weak ($r = 0.13; P = 0.05$). The

---

**Table 2** Categorical comparisons

<table>
<thead>
<tr>
<th>Variable</th>
<th>PSQI $\leq 5$</th>
<th>PSQI $&gt; 5$</th>
<th>Odds ratio (95% CI)</th>
<th>$P$-value ($\chi^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\leq 10$</td>
<td>$50$ (34.2)</td>
<td>$96$ (65.8)</td>
<td>Referent</td>
<td></td>
</tr>
<tr>
<td>$&gt; 10$</td>
<td>$19$ (21.1)</td>
<td>$71$ (78.9)</td>
<td>$1.9$ (1.1–3.6)</td>
<td>$0.03$</td>
</tr>
</tbody>
</table>

**Abbreviations**: CI, confidence interval; ESS, Epworth Sleepiness Scale; PSQI, Pittsburgh Sleep Quality Index.
distribution of ESS and PSQI scores in Figure 1 provides further illustration of the tenuous association. Examination of the relationship with AHI measurements from PSG also revealed a weak correlation ($r = -0.12$), which did not achieve statistical significance ($P = 0.07$). Participants with abnormal PSQI $> 5$ exhibited no statistically significant differences in PSG measures examined compared to those with normal PSQI scores.

In our aggregate sub-group comparison, the sex proportion and mean BMI exhibited between group differences. With the exception of the APLM index, which was relatively low frequency, the PSG measures were not statistically significantly different in these four groups categorized by normal/abnormal for ESS and PSQI.

This was a chart review study and therefore subject to the limitations inherent in a retrospective review. As the patient population was derived from a group referred for PSG, there is the risk of inclusion bias, both for men in whom sleep disorders are perhaps more readily questioned, and for those who may be more overtly symptomatic or aware of sleep problems. However, the generalizability of our findings to other populations may be questionable, although the observation of limited association between the PSQI and ESS is in keeping with the results from Buysse et al in a community-based population. These observations in both these population samples do not denigrate the value of either questionnaire assessment, but further emphasize the disparity in the facet of sleep health under examination through each instrument.

In the assessment of between-group differences or associations in terms of polysomnographic measures, we acknowledge that there is the potential for a ‘first night effect’, which would potentially impact on PSG parameters, in particular, sleep efficiency. This concern would likely be a lesser issue for objective measures such as AHI.

In summary, in this retrospective study of high-risk patients referred for PSG, there was limited association observed between these two commonly used questionnaire instruments, the ESS and the PSQI. These two questionnaires appear to evaluate different aspects of sleep. In terms of clinical application, for global assessment of patients with sleep problems, care should be taken to include instruments measuring different facets of sleep health.
Table 5  Comparison of empirically-defined PSQI/ESS subgroups

<table>
<thead>
<tr>
<th>Variable</th>
<th>PSQI ≤ 5 and ESS ≤ 10</th>
<th>PSQI ≤ 5 and ESS &gt; 10</th>
<th>PSQI &gt; 5 and ESS ≤ 10</th>
<th>PSQI &gt; 5 and ESS &gt; 10</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean BMI (kg/m²; SD)</td>
<td>32.2 (8.1)</td>
<td>34.7 (8.4)</td>
<td>33.9 (7.8)</td>
<td>36.6 (8.7)</td>
<td>0.02*</td>
</tr>
<tr>
<td>BMI (kg/m²; %)</td>
<td>30–34</td>
<td>25–29</td>
<td>35–39</td>
<td>≥40</td>
<td></td>
</tr>
<tr>
<td>Mean age (years) (SD)</td>
<td>52.6 (15.5)</td>
<td>55.3 (12.9)</td>
<td>53.3 (13.8)</td>
<td>51.8 (13.4)</td>
<td>0.79*</td>
</tr>
<tr>
<td>Sex (%)</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td></td>
</tr>
<tr>
<td>Mean SE (SD)</td>
<td>70.6 (20.1)</td>
<td>70.8 (20.1)</td>
<td>67.6 (21.7)</td>
<td>69.9 (21.3)</td>
<td>0.72*</td>
</tr>
<tr>
<td>AHI (Median, IQR)</td>
<td>17.6 (33.2)</td>
<td>28.6 (70.9)</td>
<td>12.0 (24.2)</td>
<td>18.2 (47.6)</td>
<td>0.07*</td>
</tr>
<tr>
<td>Arousing Periodic Limb Movements (APLM, Mean, IQR)</td>
<td>14.5 (16.0)</td>
<td>11.5 (13.0)</td>
<td>15.0 (16.0)</td>
<td>13.0 (15.0)</td>
<td>0.48*</td>
</tr>
<tr>
<td>Arousal Index (AI)</td>
<td>0.0 (0.6)</td>
<td>0 (0)</td>
<td>0.6 (2.3)</td>
<td>0.7 (3.2)</td>
<td>0.67*</td>
</tr>
<tr>
<td>RERA (Median, IQR)</td>
<td>0.6 (3.5)</td>
<td>0 (1.0)</td>
<td>0.6 (2.3)</td>
<td>0.7 (3.2)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *P-value based on ANOVA; ^P-value based on Kruskal-Wallis test; δP-value based on chi-square test.

Abbreviations: ANOVA, analyses of variance; AHI, Apnea–Hypopnea Index; AI, Arousal Index; APLM, Arousing Periodic Limb Movements; BMI, Body Mass Index kg/m²; ESS, Epworth Sleepiness Scale; IQR, interquartile range; PSQI, Pittsburg Sleep Quality Index; RERA, respiratory effort-related arousals; SD, standard deviation; SE, sleep efficiency.

Acknowledgments

The authors gratefully acknowledge support from University of Saskatchewan Department of Medicine Research Fund. Prosanta Mondal was also supported by the Western Regional Training Centre for Health Services Research.

Disclosure

The authors report no conflicts of interest in this work.

References


